



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/057,331	01/24/2002	Andrei Z. Broder	22136-06658	1535

7590 06/24/2005

Seth Ostrow  
BROWN RAYSMAN MILLSTEIN FELDER AND STEINER LLP  
900 Third Avenue  
New York, NY 10022-4728

EXAMINER
----------

HILLERY, NATHAN

ART UNIT	PAPER NUMBER
----------	--------------

2176

DATE MAILED: 06/24/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	Application No. 10/057,331	Applicant(s) BRODER ET AL.	
	Examiner Nathan Hillery	Art Unit 2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

#### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) ☒ Responsive to communication(s) filed on 07 April 2005.
- 2a) ☒ This action is FINAL. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) ☒ Claim(s) 1-12 and 14-16 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-12 and 14-16 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- |   |   |
|---|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892)                        | 4) <input type="checkbox"/> Interview Summary (PTO-413)                     |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)    | Paper No(s)/Mail Date. _____  |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| Paper No(s)/Mail Date _____   | 6) <input type="checkbox"/> Other: _____                                    |

158

### DETAILED ACTION

1. This action is responsive to communications: Amendment filed on 4/7/05.
2. Claims 1 – 12, and 14– 16 are pending in the case. Claims 1, 9, and 16 are independent.
3. The rejection of claims 1 – 8 under 35 U.S.C. 101 as being non-statutory has been withdrawn.
4. The rejection of claims 9 – 12, and 14 – 16 under 35 U.S.C. 101 as being non-statutory has been maintained.
5. The rejection of claims 1 – 12, and 14 – 16 under 35 U.S.C. 103(a) as being unpatentable has been maintained.

### ***Claim Rejections - 35 USC § 101***

6. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

7. Claims 9 – 12 and 14 – 16 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.
8. Claims 9 – 12, 14, and 15 recite non-functional descriptive material, specifically software on a computer readable medium not tangibly embodied to a computer. Consequently, the claimed invention does not require the technical or useful arts and, thus, fails to define patentable subject matter. The rejection to these claims may be overcome if an inference to some form of hardware is claimed or it can be shown within the specification that the machine-readable medium can not be a signal or carrier wave.

9. Claim 16 is software per se and is not tangibly embodied to a computer system. Consequently, the claimed invention does not require the technical or useful arts and, thus, fails to define patentable subject matter. The rejection to these claims may be overcome if an inference to some form of hardware is claimed.

10. Further, to expedite a complete examination of the instant application the claims rejected under 35 U.S.C. 101 (nonstatutory) above are further rejected as set forth below in anticipation of applicant amending these claims to place them within the four statutory categories of invention.

***Claim Rejections - 35 USC § 103***

11. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

12. Claims 1 – 12 and 14– 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A).

13. **Regarding independent claim 1**, Bharat et al. teach that *the set of documents can be produced by combining the set of results from a Web search engine in response to a user query (which we call the 'start-set'), with pages that either link to or are linked from the start-set documents ... The nodes in the start set are first scored according to their connectivity, and the number of terms of the query that appear as unique sub-strings in the URL of the represented documents. The score is a weighted sum of the number of directed edges to and from a node and the number of unique sub-strings of*

*the URL that match a query term* (Column 3, lines 3 – 15), which provide for **receiving a document to be processed; locating a set of documents that include hyperlinks to the document; retrieving anchortext (sub-strings in the URL) associated with at least one of the hyperlinks, and parsing the anchortext (sub-strings in the URL) into one or more tokens**. Bharat et al. also teach that *as previously disclosed, uses the relevance weights of all of the nodes to decide whether or not to eliminate a page for user consideration. For example, prune all nodes whose relevance weight is below a predetermined threshold. The threshold can be picked in a number of ways* (Column 6, lines 22 – 27), which provide that **for each token: determining a weight for the token, determining whether the weight assigned to the token exceeds a threshold token weight**. Bharat et al. do not explicitly teach **indexing**. However, it would have been obvious to one of ordinary skill in the art to be motivated to use and/or alter the invention of Bharat et al. to provide for **indexing the document under the token, if the token weight assigned to the token exceeds the threshold token weight**, since Bharat et al. do teach that *in order to help users locate Web pages of interest, a search engine 140 maintains an index 141 of Web pages in a memory, for example, disk storage* (Column 4, lines 9 – 11) and that *we provide an improved ranking method 200 that can be implemented as part of the search engine 140. Alternatively, the method 200 can be implemented by one of the clients 110, or some other computer system on the path between the search engine and the clients* (Column 4, lines 23 – 27).

14. **Regarding dependent claims 2 and 5**, Bharat et al. teach that *specifically, in step 220, we score each page  $p$  of the input set 201 to determine a value  $\text{Score}(p)$*  225.

Let  $n_p$  be the node representing page  $p$ . The score is determined by:  $\text{Score}(p) = \text{in\_degree} + 2 \times (\text{num\_query\_matches}) + \text{out\_degree}$ , where  $\text{in\_degree}$  is the number of edges pointing at node  $n_p$ ,  $\text{num\_query\_matches}$  is the number of unique sub-strings of the URL of the page  $p$  that exactly match a term in the user's query (Column 5, lines 57 – 64), which provide for **including in the index an indication of weight for each token under which each page is indexed, and that the weight of each token is based on its frequency of occurrence within the index.**

15. **Regarding dependent claims 3 and 4**, Bharat et al. teach that *next*, we assign a relevance weight to a subset of the nodes 212. The relevance weight measures the similarity between the represented page and the query topic. As stated above, the topic implied by the user is probably broader than the query itself. Thus, matching the words of the query with the page is usually not sufficient. Instead, as described in detail below, we use a subset of the pages of the start set 201 to define a broader query topic "Q", and match the pages "P" represented in the graph with the broader query topic to determine the relevance weights of the nodes 212. Our invention is motivated by the observation that not all pages represented by nodes in the  $n$ -graph 211 are equally influential in deciding the outcome of our ranking process (Column 5, lines 21 – 33).

Bharat et al. do not explicitly teach assigning the token to a particular location; however, one of ordinary skill in the art at the time of the invention would be motivated to alter the invention of Bharat et al. to provide for **assigning to the token a location within the index corresponding to part of the page being indexed that is allocated for tokens having a higher degree of importance than other tokens in the same page, and for**

**assigning to the token a location within the index that corresponds to the beginning of the page being indexed**, since the skilled artisan would want to point the user to the exact location in the page or to the beginning of the page so that the user does not have to hunt for the exact location himself or become confused and/or overwhelmed by the results and information that he is trying to interpret.

16. **Regarding dependent claims 6 and 7**, Bharat et al. teach that *because the query topic Q 245 can include a large number of terms, and because the "vocabulary" of the various pages can vary considerably, we prefer to use term frequency weighting. More specifically, we use cosine normalization in weighting both the query topic Q and the pages P because the deviation in term vector lengths is large, specifically: ... where  $w_{iq} = \text{freq}_{iq} \times \text{IDF}_i$ ,  $w_{ij} = \text{freq}_{ij} \times \text{IDF}_i$ ,  $\text{freq}_{iq}$  is the frequency of (stemmed) term i in the query topic Q,  $\text{freq}_{ij}$  is the frequency of term i in page j, and  $\text{IDF}_i$  is an estimate of the inverse document frequency (IDF) of the term i in the corpus of documents, for example, in our case, a large representative sample of Web pages (Column 7, lines 10 – 29), which provide for determining a first frequency at which the anchor text appears in the index; determining a second frequency at which each token derived from the anchor text appears in the index; and assigning a weight to the token; wherein the weight is a function of the first and second frequencies, and dividing the first frequency by the second frequency to produce a weight quotient; and multiplying the weight quotient by an anchor text count for the token.*

17. **Regarding dependent claim 8**, Bharat et al. teach that *during a connectivity analysis phase, the remaining nodes of the pruned graph are then scored according to their connectivity to determine normalized hub and authority scores for the documents. The normalized scores are used to rank the documents* (Column 3, lines 31 – 35), which provide for **determining a normalized weight for each token**.
18. **Regarding independent claim 9**, the claim incorporates substantially similar subject matter as claim 1, and is rejected along the same rationale.
19. **Regarding dependent claim 10**, the claim incorporates substantially similar subject matter as claim 2, and is rejected along the same rationale.
20. **Regarding dependent claim 11**, the claim incorporates substantially similar subject matter as claim 5, and is rejected along the same rationale.
21. **Regarding dependent claim 12**, the claim incorporates substantially similar subject matter as claim 4, and is rejected along the same rationale.
22. **Regarding dependent claim 13**, the claim incorporates substantially similar subject matter as claim 5, and is rejected along the same rationale.
23. **Regarding dependent claim 14**, the claim incorporates substantially similar subject matter as claim 6, and is rejected along the same rationale.
24. **Regarding dependent claim 15**, the claim incorporates substantially similar subject matter as claim 8, and is rejected along the same rationale.
25. **Regarding independent claim 16**, the claim incorporates substantially similar subject matter as claim 1, and is rejected along the same rationale.



***Response to Arguments***

26. Applicant's arguments filed 4/7/05 have been fully considered but they are not persuasive.

27. In response to applicant's argument regarding claims 9 – 12, 14, and 15 under 35 USC 101, it should be noted that "machine-readable medium" can be interpreted as consisting of a signal and/or a carrier wave, which are not tangible media.

28. In response to applicant's argument regarding claim 16 under 35 USC 101, it should be noted that the Office agrees with applicant that AT&T and State St say that one test for 35 USC 101 is the useful, concrete tangible test; the claim fails such test.

29. In response to Applicant's argument(s) that Bharat does not teach Anchortext, it should be noted that the Office concedes that although Bharat does not explicitly or specifically uses the word anchortext, it would have been obvious to one of ordinary skill in the art at the time of the invention to interpret that substrings of a URL are examples of tokens in anchortext. Those of ordinary skill in the art are well aware that anchortext can simply be the URL itself. Using Applicant's own example, the Office asserts that if the web site for United Airlines was simply referenced as the URL <www.UAL.com> on a web site, then <www.UAL.com> would be the Anchortext and the unique substrings {www, UAL, com} would be the tokens of the Anchortext.

***Conclusion***


30. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Nathan Hillery whose telephone number is (571) 272-4091. The examiner can normally be reached on M - F, 10:30 a.m. - 7:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather R. Herndon can be reached on (571) 272-4136. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

  
HEATHER R. HERNDON  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2100

NH